

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-345244

(43)Date of publication of application : 14.12.1999

(51)Int.Cl.

G06F 17/30

(21)Application number : 10-154721

(71)Applicant : NIPPON TELEGR & TELEPH CORP  
<NTT>

(22)Date of filing : 03.06.1998

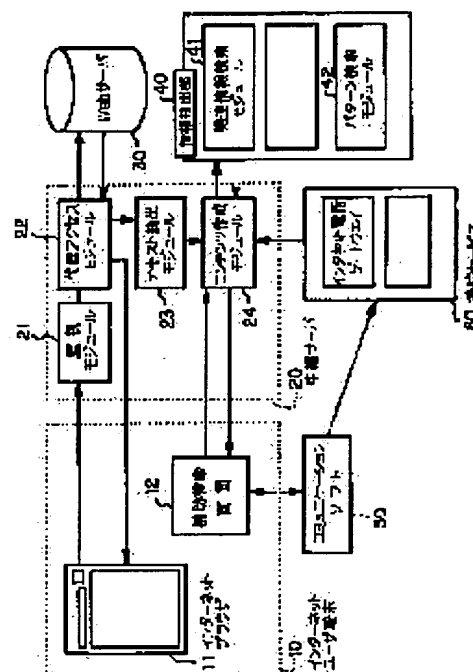
(72)Inventor : TOKUNO SATOSHI  
MIYAZAKI HAJIME  
ARAKI TAKAHIRO

## (54) METHOD AND SYSTEM FOR EXTRACTING AND DISPLAYING SIMILAR INFORMATION

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To allow an internet user to access necessary information with a simple operation by producing an auxiliary information displaying means which displays extracted auxiliary information on an auxiliary screen of a user terminal, attaching it and sending and displaying it together with response information from information providing server at the user terminal of a request source.

**SOLUTION:** A representative access module 22 accesses a Web server 30 and sends content of access results which is received from the server 30 to a user terminal. In such a case, it attaches an instruction that is to display content produced by a content producing module 24 on an auxiliary screen 12 of the user terminal to the content of the access results and sends it to a browser of the user terminal 10. A text extraction module 23 detects and extracts a part that matches with the pattern of a pattern matching table that is previously defined from the content. The module 24 produces content to be displayed on the screen 12 and sends it to the module 22.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision]

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平11-345244

(43)公開日 平成11年(1999)12月14日

(51)Int.Cl.<sup>6</sup>

G 0 6 F 17/30

識別記号

F I

G 0 6 F 15/40

15/403

3 1 0 F

3 8 0 D

審査請求 未請求 請求項の数5 O L (全 8 頁)

(21)出願番号 特願平10-154721

(22)出願日 平成10年(1998)6月3日

(71)出願人 000004226

日本電信電話株式会社

東京都千代田区大手町二丁目3番1号

(72)発明者 得能 聡

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(72)発明者 宮崎 一

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(72)発明者 荒木 隆裕

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

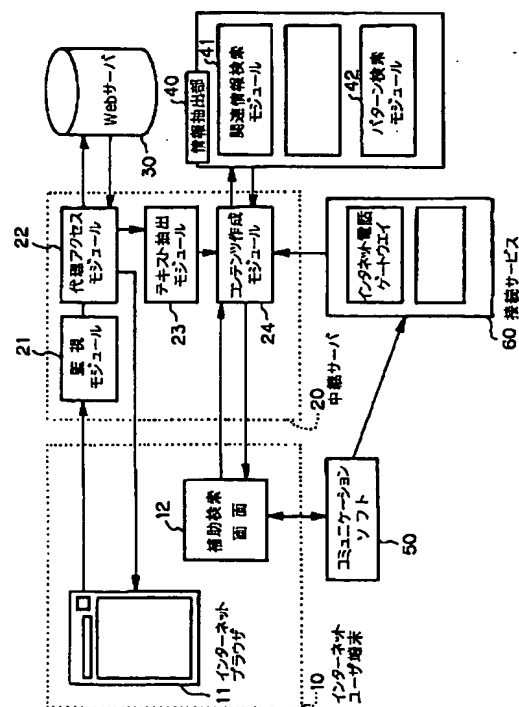
(74)代理人 弁理士 若林 忠 (外2名)

(54)【発明の名称】 類似情報抽出表示方法とシステム

(57)【要約】

【課題】 利用者の要求する検索条件を調整して必要情報を効率よく検索するとともに、ページ情報の情報提供者に対する通信手段を自動的に設定する。

【解決手段】 中継サーバ20は、利用者端末10からのアクセス要求を監視し、ユーザごとのアクセスログを保存する監視モジュール21と、Webサーバ30へのアクセスを行い、コンテンツ作成モジュール24の作成した内容を利用者端末の補助画面12に表示する命令をアクセス結果のコンテンツに付加して利用者端末のブラウザに送る代理アクセスモジュール22と、代理アクセスモジュール22が入手したコンテンツの内容からパターンマッチングテーブルのパターンにマッチする電話番号やメールアドレス等の通信情報を抽出するテキスト抽出モジュール23と、テキスト抽出モジュールの出力に自動発信用のリンクボタンを埋め込んで補助画面に表示するコンテンツを作成するコンテンツ作成モジュール24とを有する。



## 【特許請求の範囲】

【請求項 1】 利用者端末の要求により、情報提供サーバの保有する情報をネットワークの中継サーバを介して前記利用者端末の画面に表示し、前記利用者端末でブラウジングを行なう情報表示方法において、前記中継サーバが、前記利用者端末から要求された情報提供サーバにその保有する情報の提供を要求し、前記情報提供サーバの応答情報から所定の抽出条件に該当する情報を抽出して補助情報とし、前記抽出した補助情報を利用者端末の補助画面に表示する補助情報表示手段を生成し、前記補助情報表示手段を付加して前記情報提供サーバの応答情報とともに前記要求元の利用者端末に送って表示する手順を有することを特徴とする類似情報抽出表示方法。

【請求項 2】 前記情報提供サーバの応答情報から補助情報を抽出する抽出条件は、応答情報中に記載されるキーワードと、前記キーワードに続く文字列を含む請求項 1 に記載の類似情報抽出表示方法。

【請求項 3】 前記キーワードは、情報提供元の連絡先に対する通信手段の種類を含む請求項 2 に記載の類似情報抽出表示方法。

【請求項 4】 保有する情報をネットワークを介して公開する情報提供サーバと、前記公開される情報を画面に表示してブラウジングを行なう利用者端末と、前記利用者端末と情報提供サーバ間の情報転送を中継する中継サーバを含む情報表示システムにおいて、前記中継サーバは、利用者端末から情報提供サーバへのアクセス要求を監視して、発生したアクセス要求を受け付けるとともに、利用者ごとのアクセスログを保存するアクセス監視手段と、前記アクセス監視手段の受け付けたアクセス要求を情報提供サーバに転送し、情報提供サーバから送られる応答情報に、作成されたコンテンツを含む所定の情報を付加して要求元の利用者端末に転送する代理アクセス手段と、前記情報提供サーバから送られた応答情報から所定の補助情報を抽出する補助情報抽出手段と、前記抽出された補助情報にリンクボタンを埋め込んで利用者端末の補助画面に表示するコンテンツを作成し、前記代理アクセス手段に送るコンテンツ作成手段とを有することを特徴とする類似情報抽出表示システム。

【請求項 5】 前記中継サーバの代理アクセス手段が付加する所定の情報は、前記抽出した補助情報に含まれる連絡先に対して発信可能な通信手段の情報が含まれる請求項 4 に記載の類似情報抽出表示システム。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は、利用者端末の要求

により、情報提供サーバの保有する情報をネットワークの中継サーバを介して利用者端末の画面に表示し、利用者端末でブラウジングを行なう情報表示システムにおける類似情報抽出表示方法とそのシステムに関する。

## 【0002】

【従来の技術】 全世界にわたるインターネット上の情報の情報量は膨大なもので、その情報検索手段としてワールド・ワイド・ウェブ、以下 Web という、が用いられ、その情報の殆どが情報提供者である Web サーバにホームページという形のインデックスが掲載されている。

【0003】 従来の典型的な Web の情報検索表示システムを図 5 に示す。

【0004】 このシステムは、インターネットまたはイントラネットのクライアントである利用者端末 10 と、プロキシ・サーバと呼ばれる利用者端末 10 からのネットワークへの接続要求を中継する中継サーバ 20 と、Web を用いて情報提供を行っており、ネットワーク上にホームページを公開している Web サーバ 30 と、指定されたキーワードによりホームページを検索する検索エンジン 40 と、ホームページのデータベースを有するデータベース 50 と、利用者端末 10 と Web サーバ 30 の間の通信手段である電子メール、電話、FAX 等の通信網 60 とからなる。

【0005】 情報を閲覧しようとする利用者は、まず、中継サーバ 20 を指定し、Web サーバのアドレスとして URL (ユニバーサル・リソース・ロケーション) を入力して Web 情報の検索ソフトウェアであるブラウザを起動する。

【0006】 中継サーバ 20 は、指定された Web サーバ 30 に情報提供を要求し、この要求に回答して Web サーバ 30 から送られる情報を要求元の利用者端末 10 に転送する。

【0007】 利用者は、中継サーバ 20 を介して送られた Web サーバ 30 の情報をその端末の画面に表示して閲覧することができる。

【0008】 利用者がさらに関連する情報を検索しようとするときは、検索のためのキーワードを検索エンジン 40 に送り、検索エンジン 40 がこのキーワードによりデータベース 50 を検索する。そして、検索エンジン 40 により検索されたホームページが利用者端末 10 の画面に表示される。

【0009】 ホームページのコンテンツ (HTML ファイル) の中に、そのコンテンツを搭載しているサーバによってこれらのテキストがクリックされた場合の動作を規定したプログラムが予め実装されている場合は、利用者端末 10 に表示されたホームページ中に設けられるリンクボタンをクリックすることにより、そこにリンクされた他のホームページやコンテンツ情報を次々とたどって読み出すことができる。

【0010】さらに、利用者がホームページやコンテンツ情報に記載された情報内容に関する問い合わせやカタログ商品の注文等のためにホームページの掲載者とコミュニケーションを行なおうとする場合は、そのホームページに記載されている連絡先の電話番号、FAX番号、電子メールアドレス等をインターネット上、または他の通信網60に送出してホームページの掲載者と通信していた。

【0011】検索エンジンは、大量の情報をいかに能率よく分類、検索して必要なものだけを取り出すかが非常に重要で、それを実現する技術の開発が要求されている。

【0012】現在知られている主なものには、1) キーワードを入力して、予め登録されているインターネット上の場所を表すURL (ユニフォーム・リソース・ロケータ) からそのキーワードに合致するものを抽出する検索エンジンのサービスがある。また、2) エージェントと呼ばれるものをネットワーク上に送り出し、入力された商品のあるURLをさがしてくるもの、3) 予めサイトのカタログ中のアイテムに対してユーザが嗜好を入力すると、エージェントが他のユーザの統計からそのユーザの嗜好に合いそうなアイテムを検索して提供するもの等がある。

【0013】具体的には、例えば、バナー広告 (ここをクリックすると広告主のホームページにジャンプする記事中の小さな広告) の配信を個人ユーザーの行動に応じて最適化するために開発されたもので、サイト内のユーザーのあらゆるオンライン行動、例えばキーワード入力、ページ参照、ディレクトリや特定テーマの選択利用状況等、をリアルタイムに追跡してユーザーの興味分野や反応を学習し、ユーザーごとに最も効果的なバナー広告を瞬時に判断して表示する検索システムや、ドキュメントのキーワードを入力すると、そのキーワードのみならず、キーワードの持つコンセプトに密接に関連する情報まで検索を広げてドキュメントを抽出する検索エンジン、あるいは、Web上の情報のみならず、Web上にないデータベースに蓄積されたエンターテイメント情報や専門誌等からの情報 (Special Collection) も同時に検索でき、さらに、検索結果を一覧表示するだけでなく、ドメイン別にカテゴライズして表示するカスタム・サーチ・フォルダー機能も持っている検索エンジン等が発表されている。

【0014】

【発明が解決しようとする課題】 上述のように従来の検索エンジンは、検索項目が自分の入力した単語のキーワードに限定されてしまい、関連事項で興味のある情報もそのキーワードが入力されないかぎり検索できなかったり、関連事項が検索できるものでも、予めユーザのデータを入力して置かなければならないという問題があった。しかもこれらのサービスでは、必要情報を検索する

ためには、ユーザ自らが検索エンジンのサイトにアクセスしに行かなければならないという煩雑さがあった。

【0015】また、従来のインターネット電話、電子メール等の各種メディアでは、各端末にプログラムと連絡先情報が組込まれていないので、エンドユーザーの端末から直ぐに電話、FAX、電子メール等を発信することができなかった。

【0016】本発明の目的は、インターネットの利用者が必要な情報を簡単な操作で閲覧できる類似情報抽出表示方法とシステムを提供することにある。

【0017】さらに、インターネットの利用者が閲覧した情報中の連絡先に対する通信手段を自動的に設定できるようにすることにある。

【0018】

【課題を解決するための手段】 本発明の類似情報抽出表示方法は、中継サーバが、利用者端末から要求された情報提供サーバにその保有する情報の提供を要求し、情報提供サーバの応答情報から所定の抽出条件に該当する情報を抽出して補助情報とし、抽出した補助情報を利用者端末の補助画面に表示する補助情報表示手段を生成し、補助情報表示手段を付加して情報提供サーバの応答情報とともに要求元の利用者端末に送って表示する手順を有する。

【0019】また、情報提供サーバの応答情報から補助情報を抽出する抽出条件は、応答情報中に記載されるキーワードと、そのキーワードに続く文字列を含むことが望ましい。

【0020】また、そのキーワードには、少なくとも情報提供元の連絡先に対する通信手段の種類を含めるのがよい。

【0021】本発明の類似情報抽出表示システムは、利用者端末から情報提供サーバへのアクセス要求を監視して、発生したアクセス要求を受け付けるとともに、利用者ごとのアクセスログを保存するアクセス監視手段と、アクセス監視手段の受け付けたアクセス要求を情報提供サーバに転送し、情報提供サーバから送られる応答情報に、作成されたコンテンツを含む所定の補助情報を付加して要求元の利用者端末に転送する代理アクセス手段と、情報提供サーバから送られた応答情報から所定の補助情報を抽出する補助情報抽出手段と、抽出された補助情報にリンクボタンを埋め込んで利用者端末の補助画面に表示するコンテンツを作成し、代理アクセス手段に送るコンテンツ作成手段とを有する中継サーバを含む。

【0022】また、好ましくは、サーバの代理アクセス手段が付加する所定の情報は、抽出した補助情報中に含まれる連絡先に対して発信可能な通信手段の情報を含む。

【0023】

【発明の実施の形態】 次に、本発明の実施の形態について図面を参照して説明する。

【0024】図1は、本発明の1実施例のシステム構成図である。

【0025】本実施例の情報検索システムの主要部は、ブラウジングによりインターネットの情報閲覧を行なう利用者端末10と、Webの情報提供者であるWebサーバ30と、利用者端末10の情報要求を中継してWebサーバ30の提供する情報を利用者端末10に返す中継サーバ20と、情報抽出部40と、コミュニケーションソフト50と、接続サービス部60とからなる。

【0026】本実施例の中継サーバ20は、監視モジュール21と、代理アクセスモジュール22と、テキスト抽出モジュール23と、コンテンツ作成モジュール24とを有する。

【0027】監視モジュール21は、利用者端末10からのアクセス要求を監視し、代理アクセスモジュールにアクセス要求を行なうとともに、利用者ごとのアクセスログを保存する。

【0028】代理アクセスモジュール22は、Webサーバ30へのアクセスを行い、アクセス結果のWebサーバ30から受信したコンテンツを利用者端末に送る。このとき、コンテンツ作成モジュール24の作成した内容を利用者端末の補助画面12に表示する命令をアクセス結果のコンテンツに付加して利用者端末10のブラウザに送る。

【0029】テキスト抽出モジュール23は、代理アクセスモジュール22が入手したアクセス結果のコンテンツの内容に漢字コード変換等の後続の処理に対する準備処理を行なうと共に、そのコンテンツから予め定められたパターンマッチングテーブルのパターンにマッチする部分を検出して抽出する。パターンマッチングテーブルのパターンのキーワードとしては、少なくとも情報提供者の電話番号、アドレス等を含むことが望ましい。

【0030】コンテンツ作成モジュール24は、テキスト抽出モジュール23から出力されたテキストにリンクボタンを埋め込んで補助画面12に表示するコンテンツを作成し、代理アクセスモジュール22に送る。

【0031】情報抽出部40は、関連情報検索モジュール41と、パターン検索モジュール42とを有する。情報抽出部40には、これらの外、ホームページ上のいわゆるバナー広告などを表示する固定情報表示手段等も備えることができる。

【0032】関連情報検索モジュール41は、キーワードではなく、利用者がどのホームページにアクセスしたかというブラウジング行為そのものを検索条件として、そのホームページと関連性や類似性のある他のホームページをも表示する。

【0033】パターン検索モジュール42は、Web30から受信した情報中に、電話番号、FAX番号、電子メールアドレス等の予め定められたパターンの情報があるか検索する。

【0034】そして、代理アクセスモジュール22から送られる補助情報中にこれらのパターンを識別できるテキストが含まれている場合は、これらテキストがクリック可能な属性に変更されて補助画面12に表示されるので、これらの電話番号、FAX番号、電子メールアドレス等をクリックすると、電話、FAX、電子メール等を発信するためのプログラムが中継サーバから利用者端末10にダウンロードされる。

【0035】コミュニケーション・ソフト50は、補助画面12に表示された電話番号、FAX番号、電子メールアドレス等がクリックされたとき、その宛先に電話、FAX、電子メール等を発信するためのプログラムである。

【0036】接続サービス部60は、インターネット電話ゲートウェイ51と、MCW52等のWeb以外の通信網である。

【0037】次に、本実施例の動作について図面を参照して説明する。

【0038】図2及び図3は、本発明の中継サーバの主要な動作の1例を示す内部シーケンス図である。

【0039】利用者端末がキーワードを指定してWebサーバへのアクセス要求のハイパーテキスト転送プロトコル(HTTP)を中継サーバに送ると、中継サーバは、代理アクセスモジュールによりそのHTTP要求を変換してWebサーバ30に転送する。

【0040】Webサーバ30は、中継サーバ20からHTTP要求を受信すると、この要求に対する応答をHTMLで中継サーバ20に返す。

【0041】中継サーバ20は、テキスト抽出モジュール23によりWebサーバ30からのHTML応答をテキストに変換し、指定されたキーワードが検出されると、テキストをデータフォーマットHTML、漢字コードSJISでコンテンツ作成モジュール24に渡す。

【0042】コンテンツ作成モジュール24は、関連情報検索モジュール41及び固定情報表示モジュール42により、テンプレートファイルを参照して類似情報のコンテンツを作成するとともに、HTML、シフトJISの漢字コードのフォーマットのデータをパターン検索モジュール43に送る。

【0043】パターン検索モジュール43は、後述のように、起動時に読み込んだパターンマッチングテーブルを参照して、電話番号、FAX番号、Eメール等の通信の宛先となるキーワードを検索し、さらに、これらの宛先にそれぞれの通信システムを介して発信するためのプログラム(Vifファイル等)を作成してコンテンツ作成モジュールに渡す。

【0044】コンテンツ作成モジュールは、補助画面表示用コンテンツにVifファイル等のプログラムへのリンクを付加してHTMLのフォーマットで利用者端末10に送る。

【0045】利用者は、端末の補助画面12の表示をクリックすることにより、中継サーバ20にプログラム転送要求が送られ、コンテンツ作成モジュール24で作成されたプログラムファイルが利用者端末10にダウンロードされるので、このプログラムを起動してインターネット・ゲートウェイ等の通信システムを介して容易に希望の宛先に発信することができる。

【0046】次に、パターン検索モジュールの動作について図4により詳細に説明する。

【0047】先ず、Webサーバからの応答情報である標準入力を読み込む（HTML）（ステップS10）。

【0048】次に、キーワード検索プロセスを起動して、電話番号、電子メールアドレス等を認識させるためのキーワードが記録されたキーワードファイルを読み込み、応答情報中のキーワードを検索する。このとき、キーワードが見つからなければ何もせずに（ステップS170）の標準出力へ進む（ステップS20）。

【0049】ここで、キーワードファイルのキーワードは、例えば、電話、電話番号、TEL、Tel、tel、PHONE、Phone、phone、ダイヤルイン等の文字列のキーワードを「電話」とし、E-mail、E-Mail、メール、メールアドレス等の文字列のキーワードは「メールアドレス」とする。

【0050】次に、応答情報中に検出されたキーワードの種類が「電話」かどうかを判定する（ステップS30）。

【0051】「電話」のキーワードであれば、キーワードの後に続く半角、全角の数字、「-」、スペースの電話番号抽出文字をそれがなくなるまで1バイトずつ読み込み（ステップS50、S60）、これらの文字が見つからなくなった場所を最後の文字とする（ステップS70、S80）。

【0052】そして、読み込んだ文字列が9バイト、または10バイトであれば、この文字列が電話番号であると見做してリストに記録し、それ以外であれば、電話番号でないと見做して処理を終了する（ステップS90、S100）。

【0053】キーワードの種類がメールアドレスであれば（ステップS110）、キーワードに続く半角の大小の英字、数字、「-」のメールアドレス抽出文字を、半角のスペース文字が見つからなくなるまで1バイトずつ読み込む（ステップS120、S130）。

【0054】半角スペース文字が見つからなくなった場所をメールアドレス抽出開始文字とし、メールアドレス抽出開始文字の後方の「@」を検索する。「@」が見つかったと、その次の文字からメールアドレス抽出文字が見つからなくなるまで1バイトずつ走査し、メールアドレス抽出文字が見つからなくなった場所をメールアドレスの最後の文字と見做す。

【0055】次に、「@」の前の文字から前方に向って

走査し、メールアドレス抽出文字が見つからなくなるまで1バイトずつ移動する。そして、メールアドレス抽出文字が見つからなくなった場所をメールアドレスの先頭の文字と見做す。

【0056】メールアドレスの先頭の文字と最後の文字に囲まれた文字列を最初の抽出文字列とし、この最初の抽出文字列中の「@」の前後にそれぞれ1バイト以上の文字が含まれる場合は、この文字列をメールアドレスと見做してリストに追加する。それ以外の場合は、メールアドレスではないとして処理を終了する。

【0057】これらの補助情報は、Web30からの応答情報中に抽出されるべき情報が検出されない場合は、利用者端末10にはWeb30からの応答情報のみが送られるのはいうまでもない。

【0058】

【発明の効果】上述のように本発明は、中継サーバが情報提供サイトの応答情報から所定の補助情報を抽出し、応答情報を利用者端末に送るとき、補助情報を利用者端末に表示させるプログラムを付加して送付して、補助情報を補助画面に表示することにより、利用者は必要な情報のキーワード等を送らなくとも簡単な操作で必要な情報を閲覧できる効果がある。

【0059】また、情報提供サイトの応答情報から当該情報中に記載された通信手段の番号及びアドレス情報が含まれている場合は、利用者端末からの発信手段を組み込むことにより、利用者から情報提供者に対する通信手段を自動的に設定できる効果がある。

【図面の簡単な説明】

【図1】本発明の1実施例のシステム構成図である。

【図2】中継サーバの動作シーケンス図の1例である。

【図3】各サイト間の情報交換のシーケンス図である。

【図4】パターン検索モジュールの動作例を示す流れ図である。

【図5】従来の典型的な情報検索システムの構成図である。

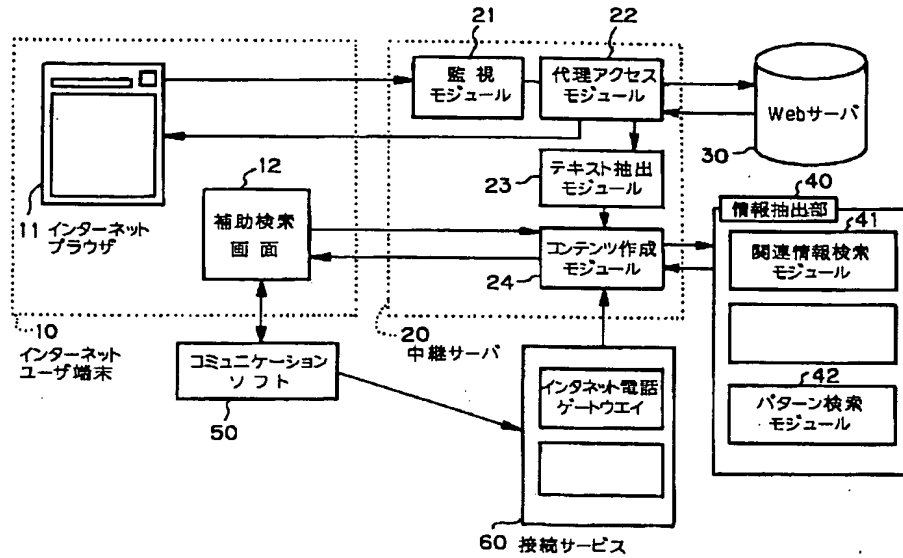
【符号の説明】

- 10 利用者端末
- 11 インターネット・ブラウザ
- 12 補助画面
- 20 中継サーバ
- 21 監視モジュール
- 22 代理アクセスモジュール
- 23 テキスト抽出モジュール
- 24 コンテンツ作成モジュール
- 30 Webサーバ
- 40 情報抽出部
- 41 関連情報検索モジュール
- 42 固定情報表示モジュール
- 43 パターン検索モジュール

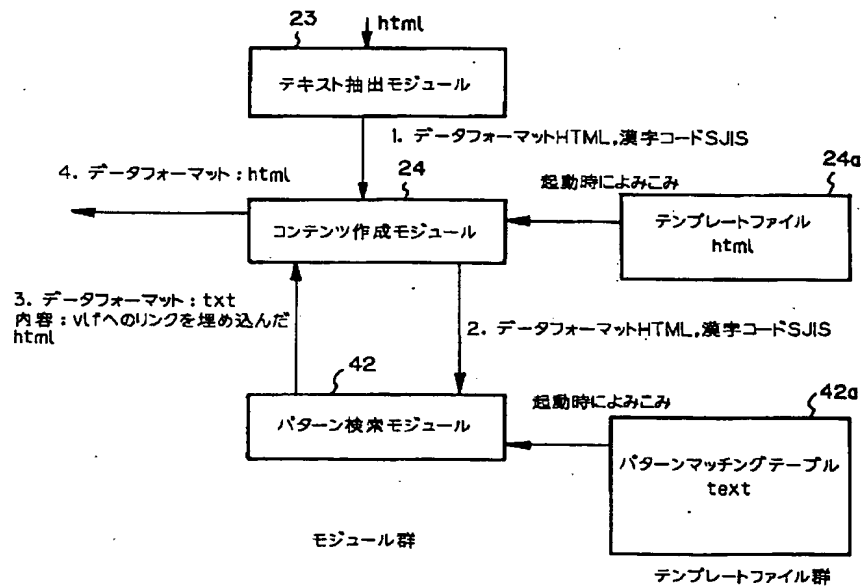
50 接続サービス部  
51 インターネット電話ゲートウェイ  
52 MCW

60 コミュニケーション・ソフト  
S10～S170 ステップ

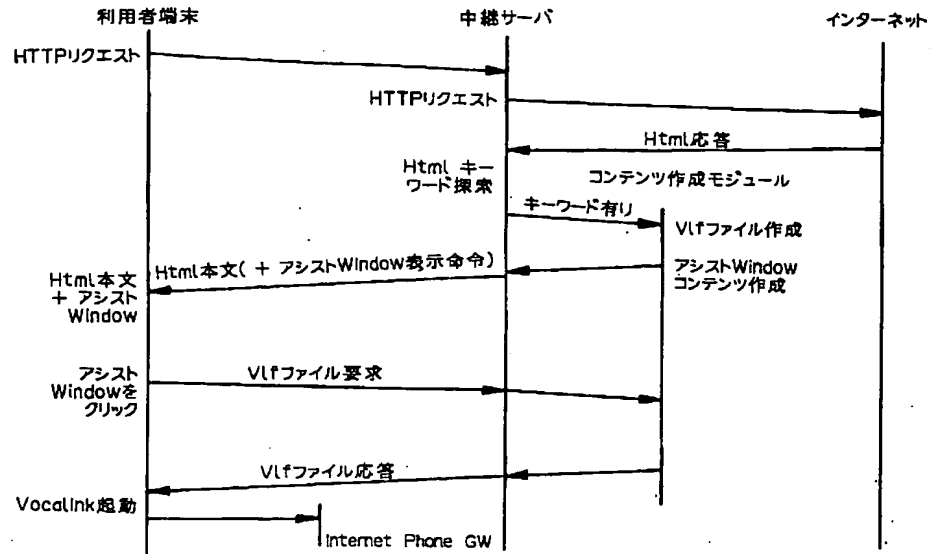
【図1】



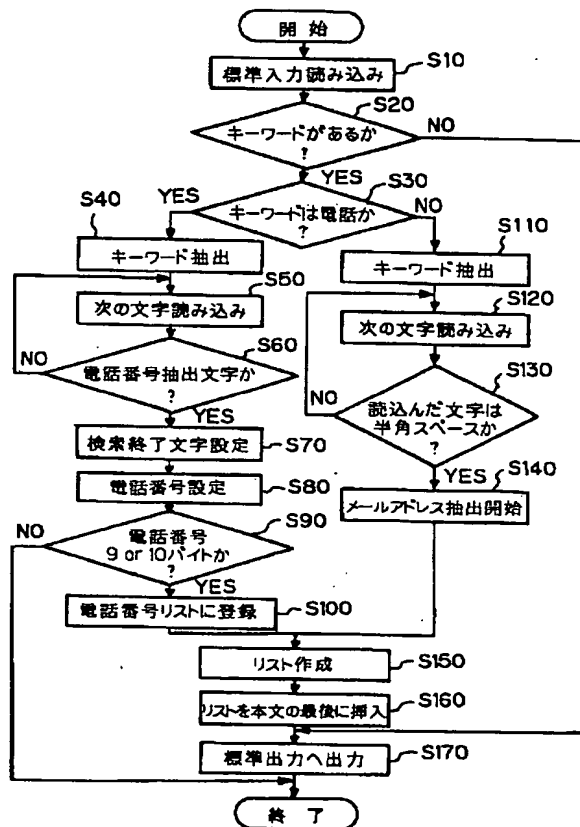
【図2】



【図 3】



【図 4】





【図 5】

